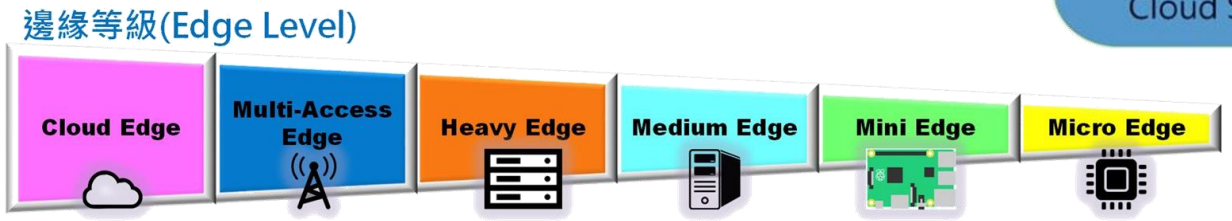
國立臺灣科技大學
NATIONAL TAIWAN UNIVERSITY OF SCIENCE AND TECHNOLOGY

Department of Computer Science & Information Engineering
資 訊 工 程 系

人工智慧與邊緣運算實務

7.5 邊緣智慧案例實作【姿態估測】



資訊工程系　許哲豪 助理教授

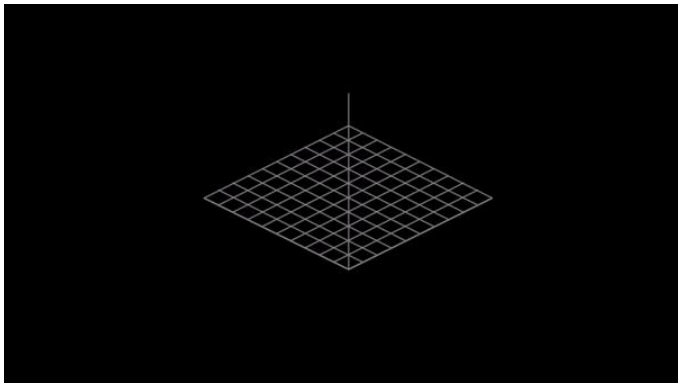
7.5 姿態估測

- 基本介紹
- OpenPose
- OpenVINO
- MediaPipe

姿態估測(Pose Estimation)

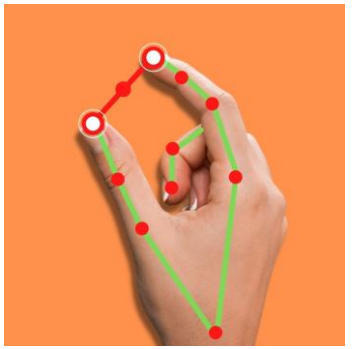
- **主要部位**
  - 頭部
  - 注視點
  - 手部
  - 全身

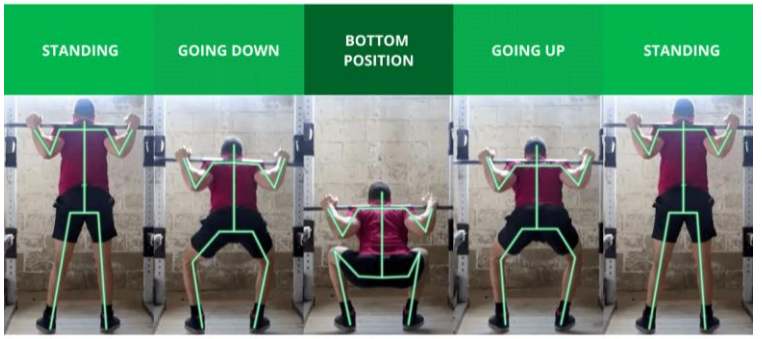
- **估測方式**
  - 2D（照片、攝影機）
  - 3D（深度攝影機）
  - 2.5D（假3D，輸入2D，輸出3D，以平面估測立體深度）

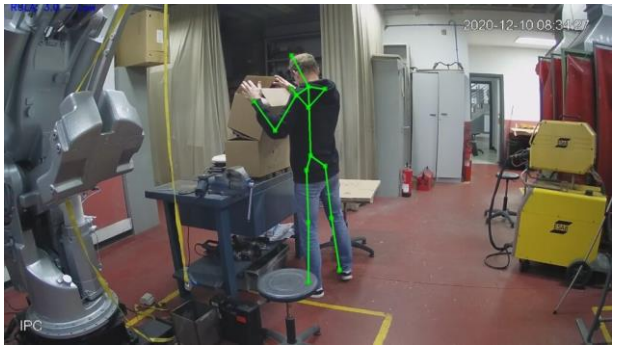


姿態估測常見用途

- 常見用途
  - 人機互動
  - 運動分析
  - 組裝確認
  - 智慧零售
  - 智慧安防
  - 智慧教室
  - 安全駕駛
  - …



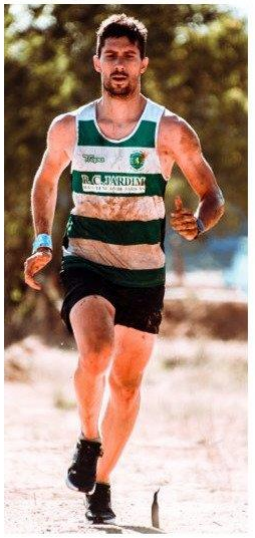




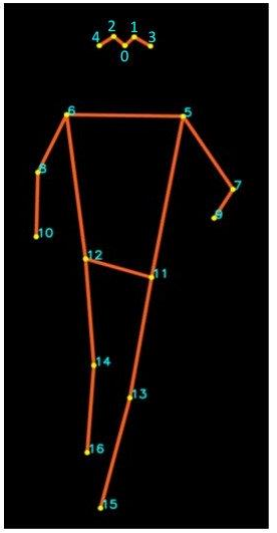
姿態估測（人體關鍵點）資料集

- 常見的2D關鍵點開放資料集
  - 微軟COCO (17點)
  - CMU OpenPose (18/25點)
    - 人體骨架 (18/25點)
    - 臉部動作 (70點)
    - 手部動作 (22點)
  - MPII (16點)
  - AI Challenge (14點)
  - LSP (14點)
  - FLIC (9點)

MS COCO關鍵點資料集



Index	Key point
0	Nose
1	Left-eye
2	Right-eye
3	Left-ear
4	Right-ear
5	Left-shoulder
6	Right-shoulder
7	Left-elbow
8	Right-elbow
9	Left-wrist
10	Right-wrist
11	Left-hip
12	Right-hip
13	Left-knee
14	Right-knee
15	Left-ankle
16	Right-ankle



Person_keypoints_val2017 範例 (JSON)

```
annotation{
    "keypoints": [x1,y1,v1,...],
    "num_keypoints": int,
    "id": int,
    "image_id": int,
    "category_id": int,
    "segmentation": RLE or [polygon],
    "area": float,
    "bbox": [x,y,width,height],
    "iscrowd": 0 or 1,
}
```

v=0 未標註, xy須為0
v=1 已標註但不可見
v=2 標註且可見

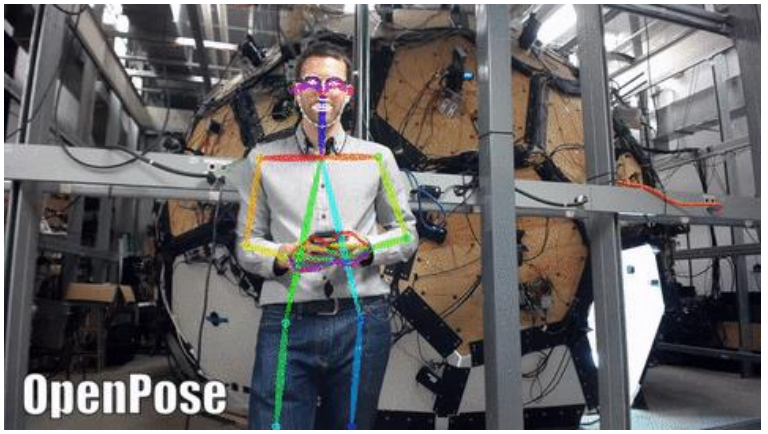
Categories字串

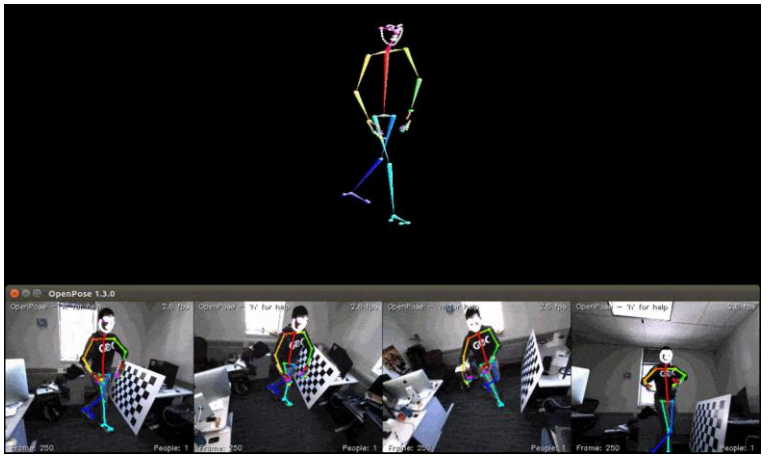
```
{
    "id": int,
    "name": str,
    "supercategory": str,
    "keypoints": [str],
    "skeleton": [edge]
}
```

關鍵點名稱 (17)
骨架連結

資料來源：https://cocodataset.org/#keypoints-2020, https://cocodataset.org/#keypoints-eval

OpenPose 基本介紹

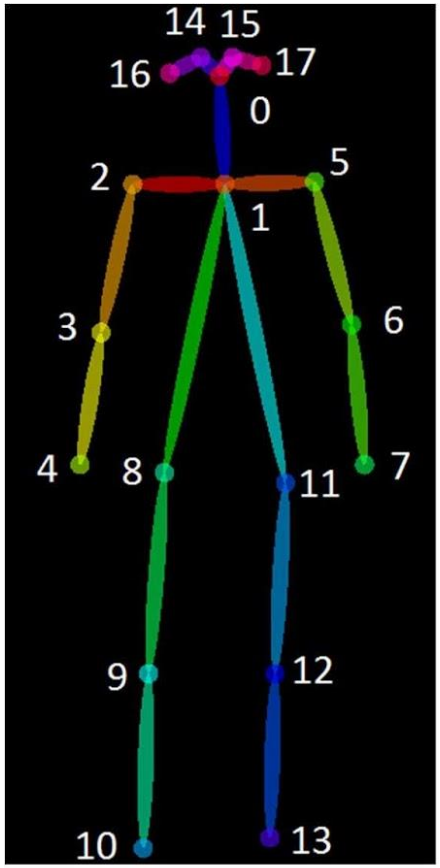




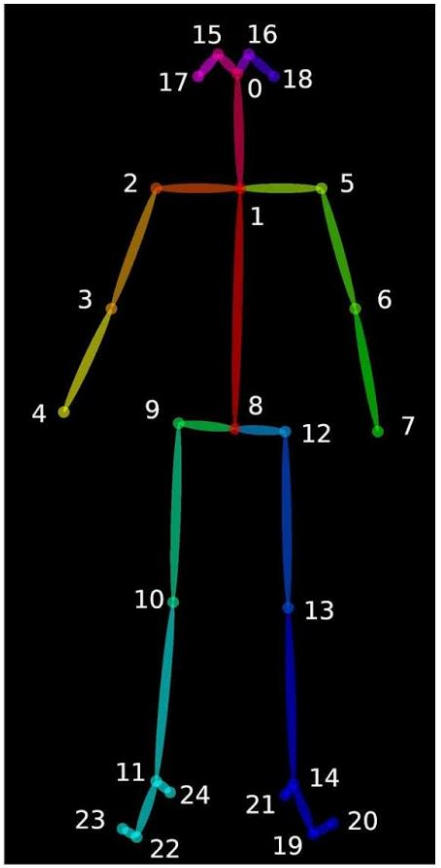
- 卡內基美濃大學(CMU) 2017年提出。
- 開源人體姿態專案，包含即時人體動作(2D/3D)、面部表情及手指運動偵測等。
- 支援Windows, Mac OS, Linux(Ubuntu)。
- 支援CPU及GPU(CUDA, OpenCL)加速。

https://github.com/CMU-Perceptual-Computing-Lab/openpose

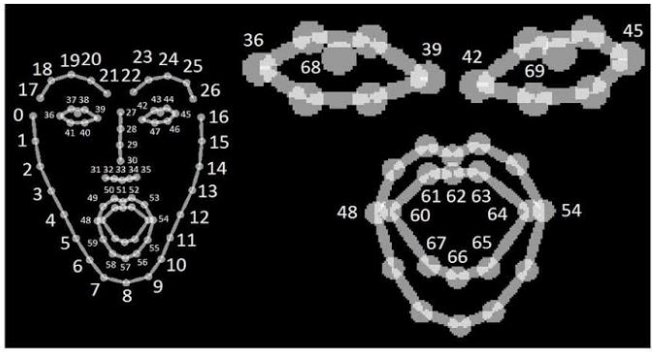
OpenPose 關鍵點定義



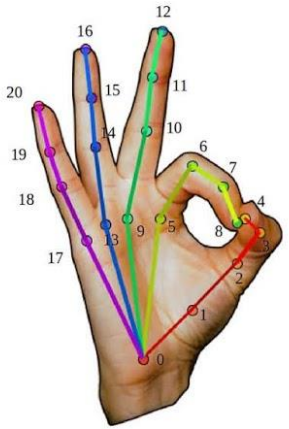
(a)



(b)



(b)



(d)

OpenPose支援輸出格式，(a)類COCO 18點，
(b) BODY 25點，(c) 臉部70點，(d)手部22點。

姿態關鍵點轉換

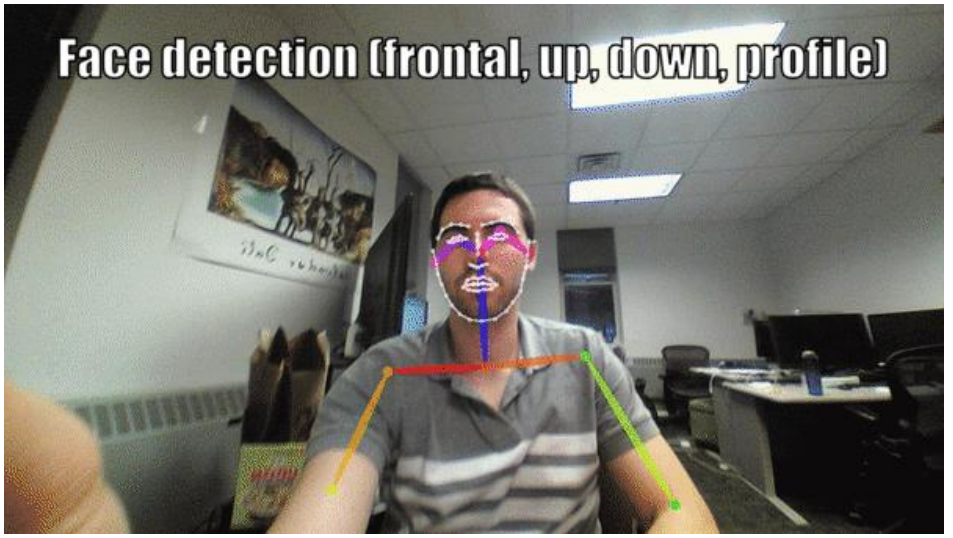
MS COCO: (17點)

{0:鼻子, 1:左眼, 2:右眼, 3:左耳, 4:右耳, 5:左肩, 6:右肩, 7:左肘, 8:右肘, 9:左腕, 10:右腕, 11:左臀（腰）, 12:右臀（腰）, 13:左膝, 14:右膝, 15:左踝, 16:右踝 }

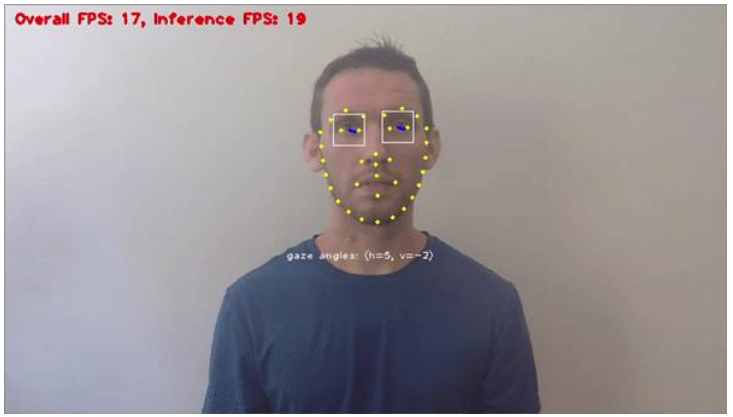
CMU OpenPose: (18點)

{0:鼻子, 1:脖子, 2:右肩, 3:右肘, 4:右腕, 5:左肩, 6:左肘, 7:左腕, 8:右臀（腰）, 9:右膝, 10:右踝, 11:左臀（腰）, 12:左膝, 13:左踝, 14:左眼, 15:右眼, 16:左耳, 17:右耳 }

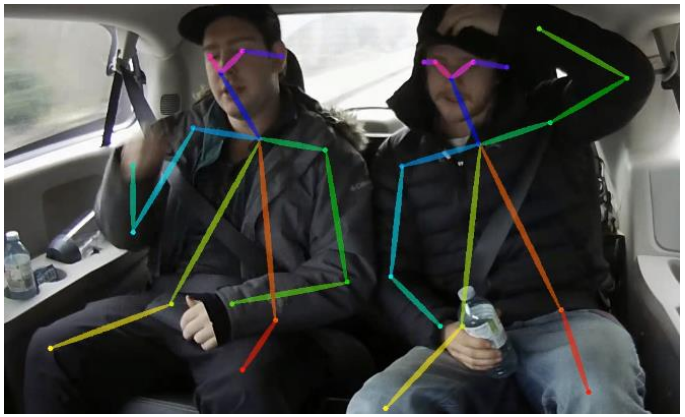
OpenPose 結果展示



OpenVINO 姿態估測範例

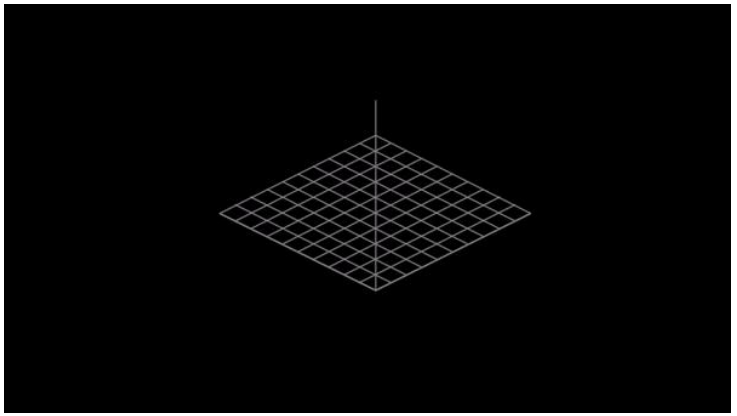


Head (Gaze) Pose Estimation



Human Pose Estimation(openpose)

Human Pose Estimation (COCO)



3D Human Pose Estimation

OpenVINO Intel's姿態估測範例

Model Name	Size	GFlops	MPara.	Precision
human-pose-estimation-0001	256x456	15.435	4.099	AP = 42.8%
human-pose-estimation-0005	288x288	5.921	8.151	AP = 45.6%
human-pose-estimation-0006	352x352	8.844	8.151	AP = 51.1%
human-pose-estimation-0007	448x448	14.325	8.156	AP = 54.3%

適用範例程式為：
Human Pose Estimation C++ Demo
Human Pose Estimation Python* Demo
　architecture_type = 0001: openpose, 0005/0006/0007: ae
Multi-Channel Human Pose Estimation C++ Demo（只適用0001）

資料來源：https://docs.openvino.ai/latest/omz_models_group_intel.html#human-pose-estimation-models

OpenVINO Public姿態估測範例

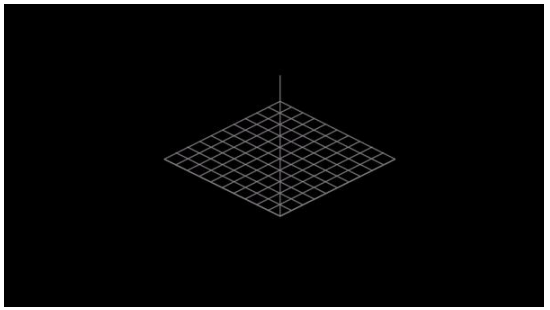
Model Name	Size	GFlops	MPara.	Precision
human-pose-estimation-3d-0001	256x448	18.998	5.074	MPJPE(mm) = 100.45
single-human-pose-estimation-0001	384x228	60.125	33.165	AP = 69.04%
higher-hrnet-w32-human-pose-estimation	512x512	92.836	28.618	AP = 64.64%

適用範例程式分別對應為：
3D Human Pose Estimation Python* Demo
Single Human Pose Estimation Demo
Human Pose Estimation Python* Demo (architecture_type = higherhrnet)

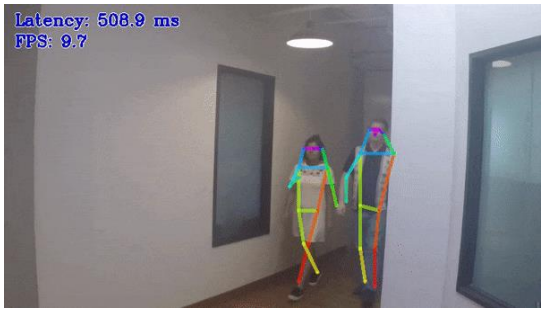
資料來源：https://docs.openvino.ai/latest/omz_models_group_public.html#human-pose-estimation-models

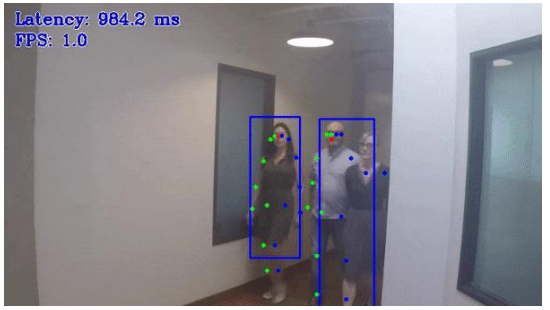
OpenVINO Demos姿態估測範例

3D (2.5D) Human Pose Estimation Python* Demo



Human Pose Estimation Python* Demo





Single Human Pose Estimation Demo (top-down pipeline)

資料來源：https://docs.openvino.ai/latest/omz_demos.html

OpenVINO Notes姿態估測範例

- Live Human Pose Estimation with OpenVINO
- 使用預訓練模型 **human-pose-estimation-0001**
- 可使用網路攝影機連續取像推論。

資料來源：https://docs.openvino.ai/latest/notebooks/402-pose-estimation-with-output.html

OpenVINO姿態估測實驗結果

human-pose-estimation-0001 @ OpenVINO 2019.2.242, i7-8950, HD630, NCS

	kinect_dance1	kinect_dance2	kinect_dance3	kinect_dance4	kinect_dance5
					
影格數量	81	168	44	111	170
首張速度(fps)	23.69 fps	24.85 fps	23.08 fps	24.32 fps	23.84 fps
最小速度(fps)	23.47 fps	24.22 fps	22.52 fps	23.62 fps	23.14 fps
最大速度(fps)	26.86fps	27.29 fps	25.85 fps	26.97 fps	27.21 fps


(a) CPU FP16 **6.81** fps


(e) GPU FP16 **23.69** fps


(b) CPU FP32 **7.06** fps


(f) GPU FP32 **15.55** fps


(c) CPU INT8 **14.01** fps


(g) GPU INT8 **14.65** fps


(d) NSC1 FP16 **2.33** fps


(h) NCS2 FP26 **4.44** fps

資料來源：https://omnixri.blogspot.com/2019/09/openvino_20.html

姿態估測指標

- Object Keypoint Similarity **(OKS)** 關鍵點相似度 (COCO)

$$OKS_p = \frac{\sum_i exp\{-d_{pi}^2/2S_p^2\sigma_i^2\}\delta(v_{pi} > 0)}{\sum_i \delta(v_{pi} > 0)}$$

- Average Precision **(AP)** 平均精準度

(單人、多人)

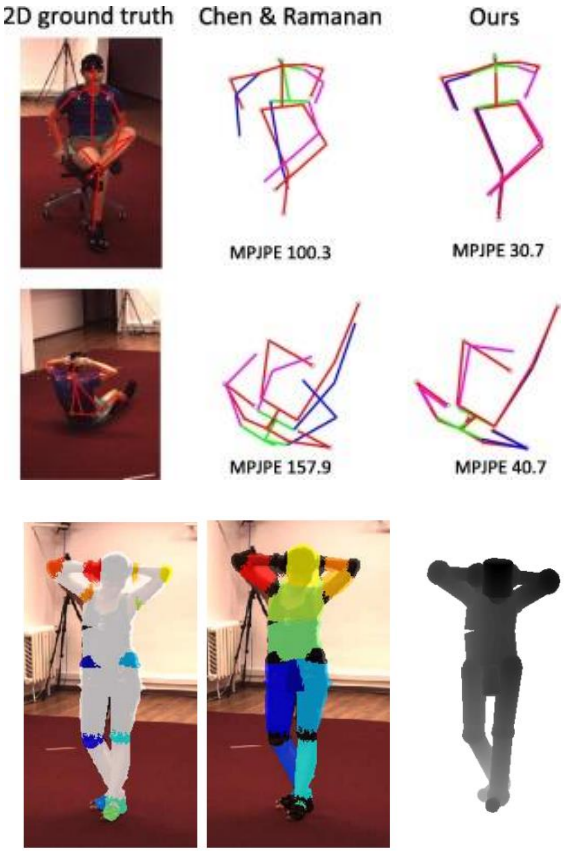
$$AP = \frac{\sum_p \delta(oks_p > T)}{\sum_p 1}$$

- Mean Per Joint Position Error **(MPJPE)**

每個關節位置誤差。

$$E_{MPJPE}(f, \mathcal{S}) = \frac{1}{N_{\mathcal{S}}} \sum_{i=1}^{N_{\mathcal{S}}} \|m_{\mathbf{f},\mathcal{S}}^{(f)}(i) - m_{\mathbf{gt},\mathcal{S}}^{(f)}(i))\|_2$$

資料來源：http://vision.imar.ro/human3.6m/pami-h36m.pdf



範例7-5-1 OpenVINO姿態估測

- 使用OpenVINO及預訓練模型human-pose-estimation-0007進行姿態估測。

- 直接以Google Colab開啟範例，可點擊下列連結：

https://colab.research.google.com/github/OmniXRI/NTUST_EdgeAI_2022/blob/main/Ch7_Implementations/Ch7-5_Pose_Estimation/Ch7-5-1_OpenVINO_Pose_Estimation/Colab_OpenVINO_Pose_Estimation.ipynb

MediaPipe 主要特色

MediaPipe

Google跨平台人臉屬性、姿態估測開源工具
CPU就能跑的很順

CPU, GPU
AI 模型加速計算

End-to-End acceleration: *Built-in fast ML inference and processing accelerated even on common hardware*

Android, iOS,
C++, Python, JS

Build once, deploy anywhere: *Unified solution works across Android, iOS, desktop/cloud, web and IoT*

多種解決方案

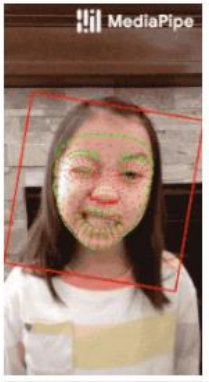
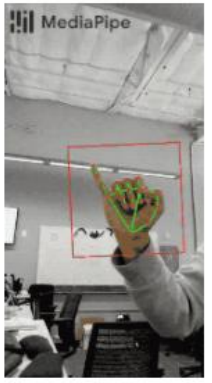
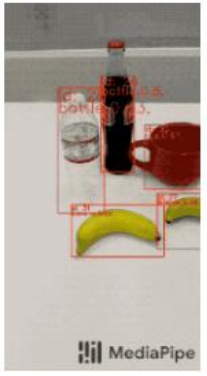
Ready-to-use solutions: *Cutting-edge ML solutions demonstrating full power of the framework*

免費及開源

Free and open source: *Framework and solutions both under Apache 2.0, fully extensible and customizable*

https://google.github.io/mediapipe/

MediaPipe 解決方案

Face Detection	Face Mesh	Iris	Hands	Pose	Holistic
					
Hair Segmentation	Object Detection	Box Tracking	Instant Motion Tracking	Objectron	KNIFT
					

MediaPipe 跨平台功能

人臉偵測
面部網點

手指骨架
人體姿態(骨架)
整體(面手肢體)
人體分割

物件盒偵測

	Android	iOS	C++	Python	JS
Face Detection	✅	✅	✅	✅	✅
Face Mesh	✅	✅	✅	✅	✅
Iris	✅	✅	✅		
Hands	✅	✅	✅	✅	✅
Pose	✅	✅	✅	✅	✅
Holistic	✅	✅	✅	✅	✅
Selfie Segmentation	✅	✅	✅	✅	✅
Hair Segmentation	✅		✅		
Object Detection	✅	✅	✅		
Box Tracking	✅	✅	✅		
Instant Motion Tracking	✅				
Objectron	✅		✅	✅	✅
KNIFT	✅				

適用Python
Google Colab

MediaPipe Colab範例程式

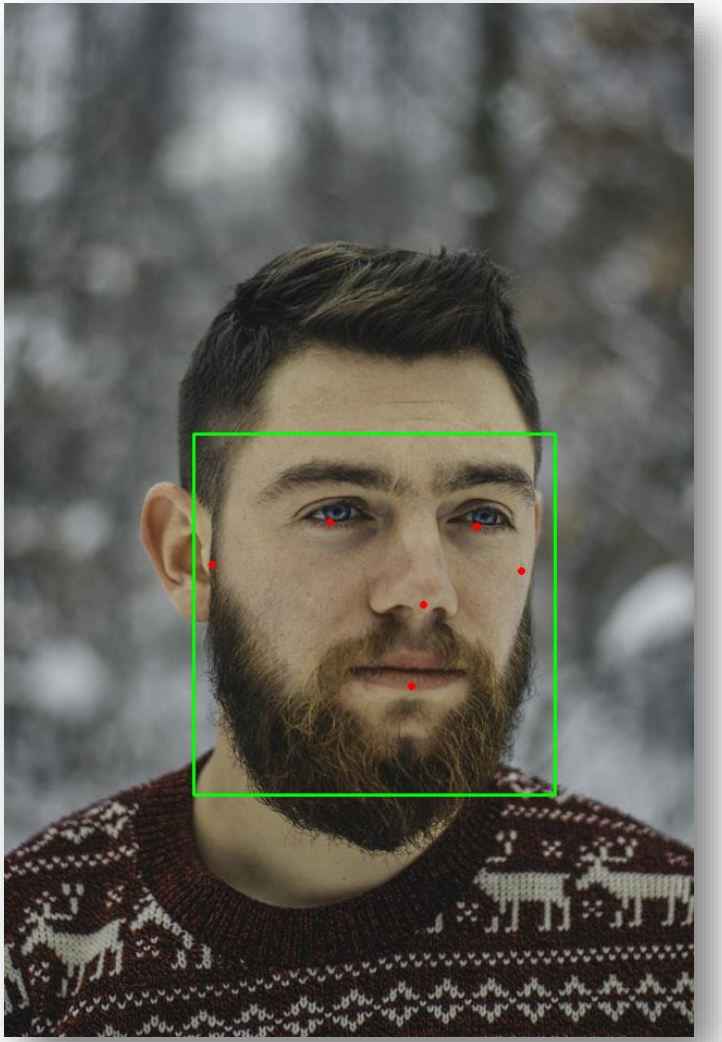
Getting Started / MediaPipe in Python / on Colab

- MediaPipe Face Detection Colab
- MediaPipe Face Mesh Colab
- MediaPipe Hands Colab
- MediaPipe Holistic Colab
- MediaPipe Objectron Colab
- MediaPipe Pose Colab
- MediaPipe Selfie Segmentation Colab

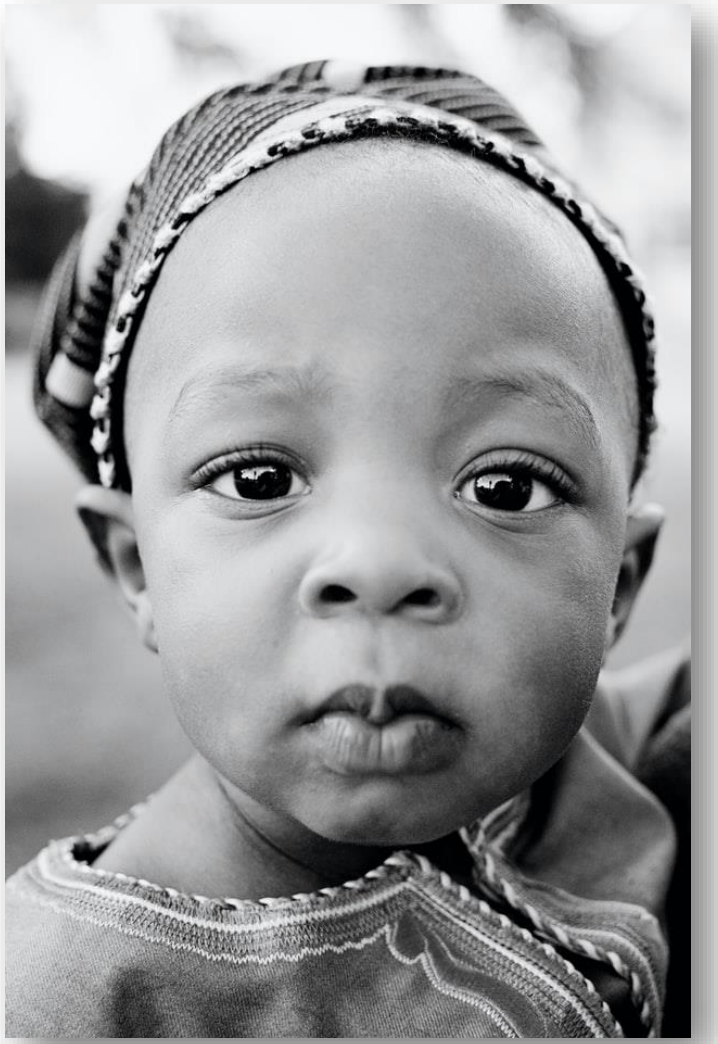
- 要先登入Google帳號才能使用
- 勿使用私隱模式或關閉Cookie下開啟網頁視窗

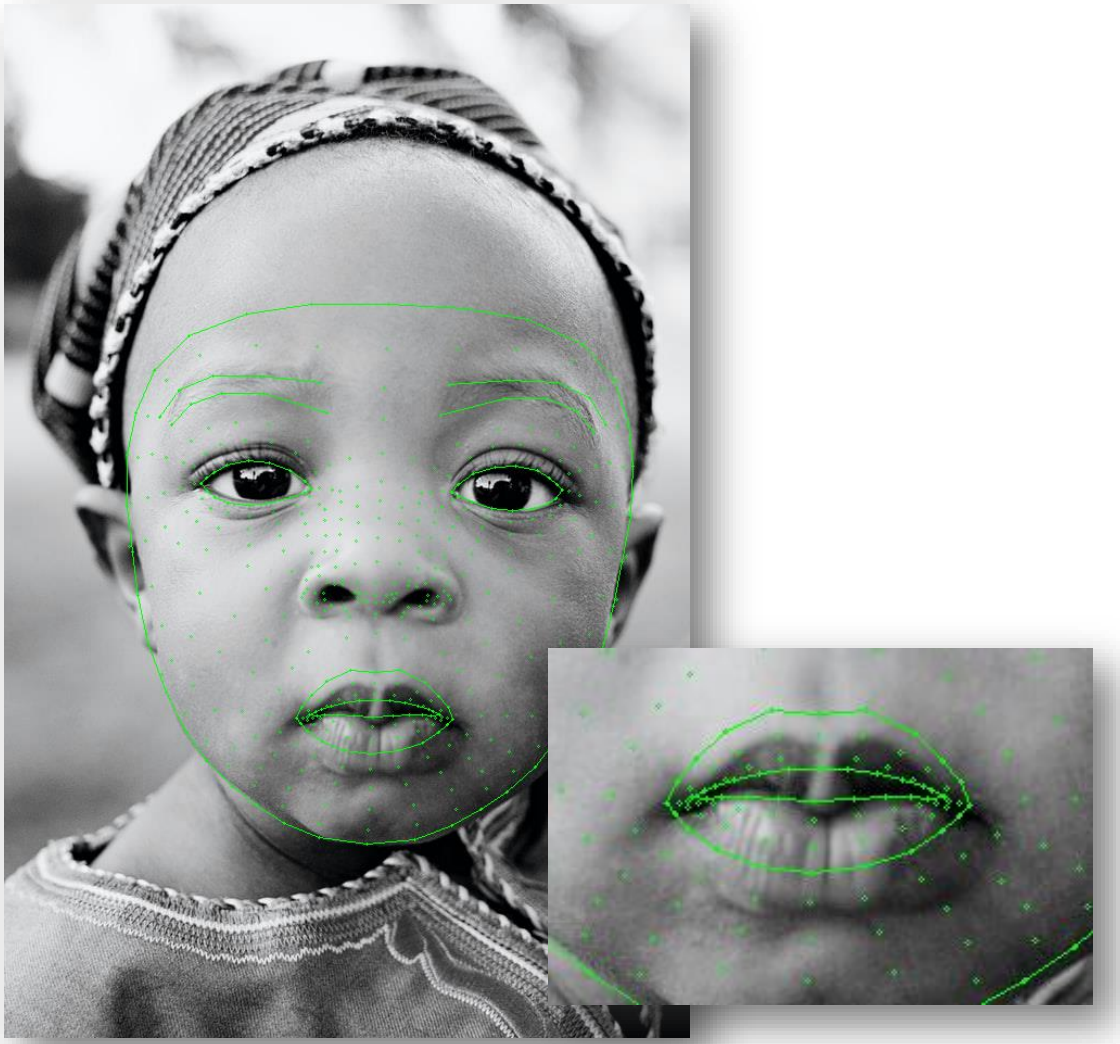
https://google.github.io/mediapipe/getting_started/python.html

MediaPipe 人臉偵測

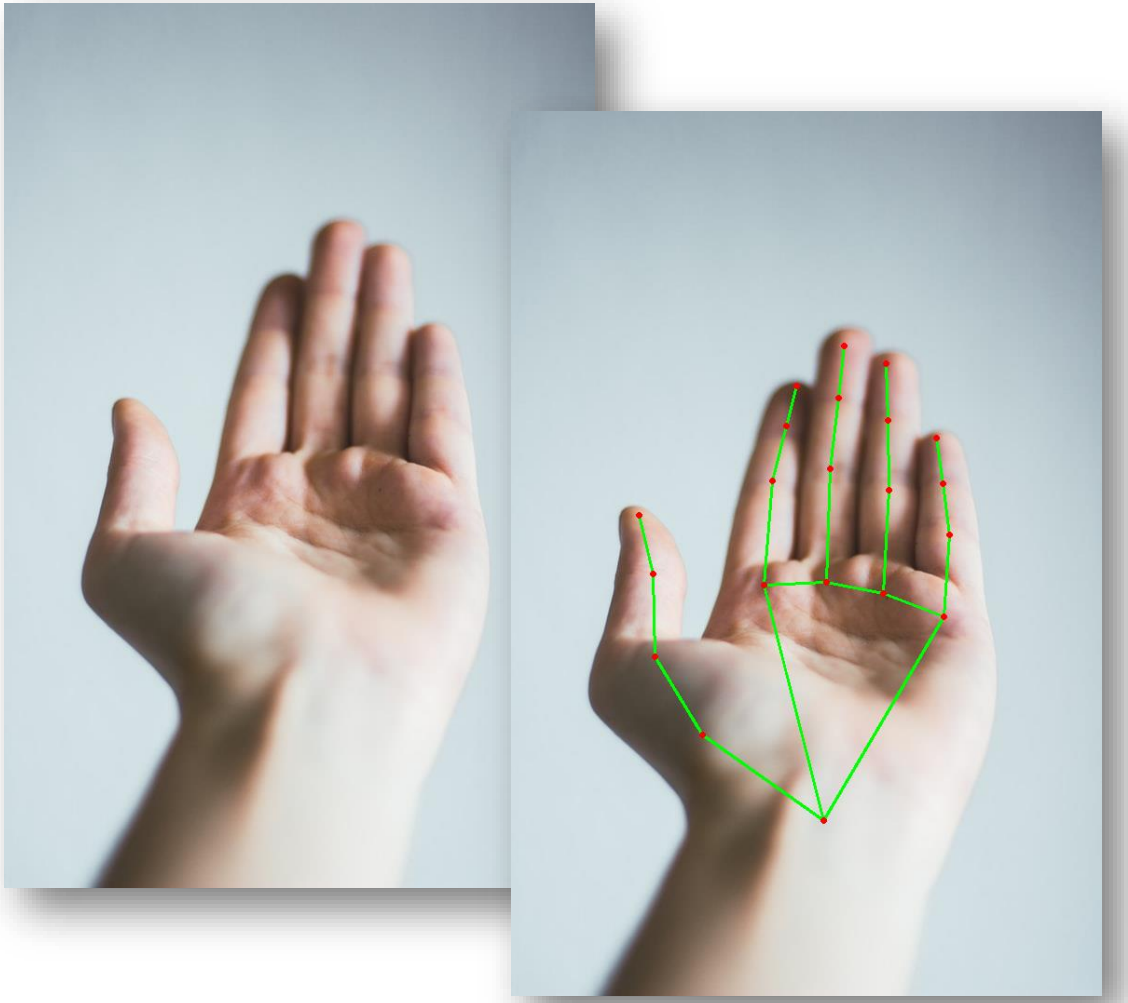


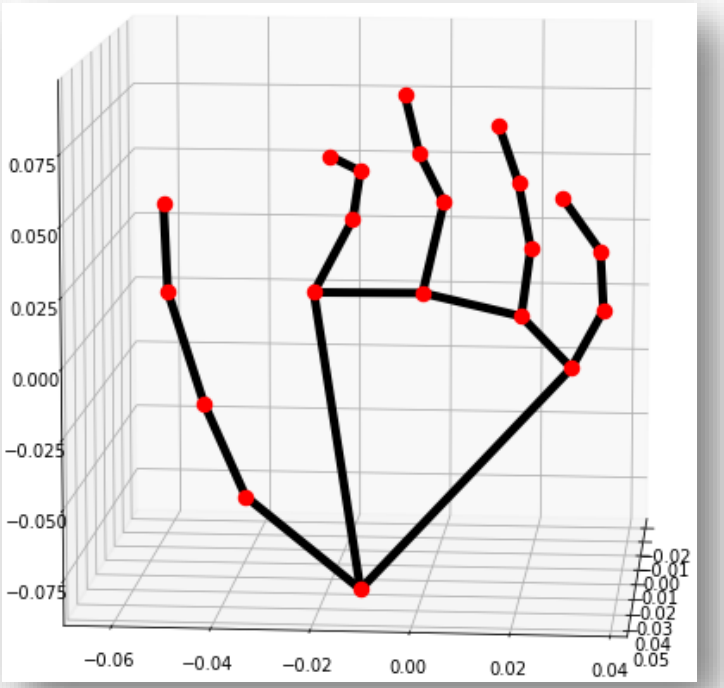
MediaPipe 面部網點



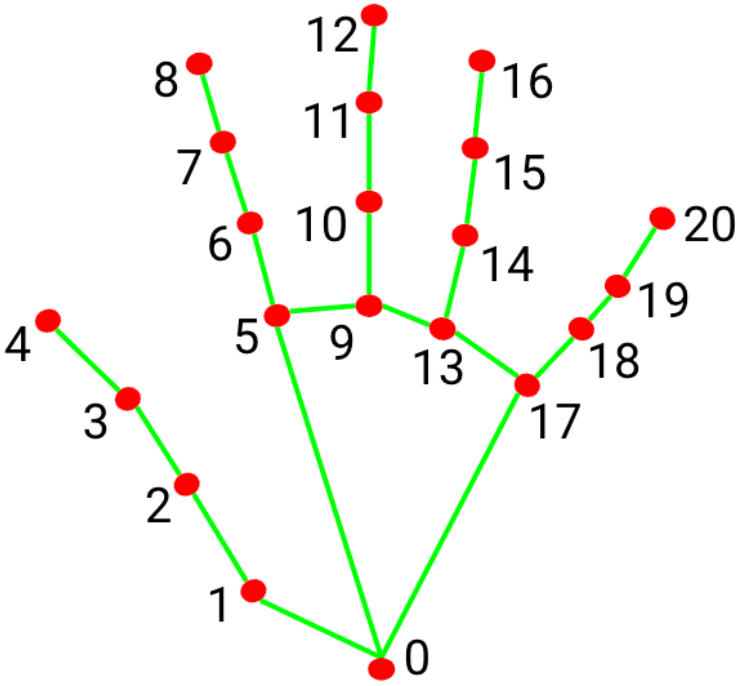


MediaPipe 手指骨架





MediaPipe 手指關鍵點定義

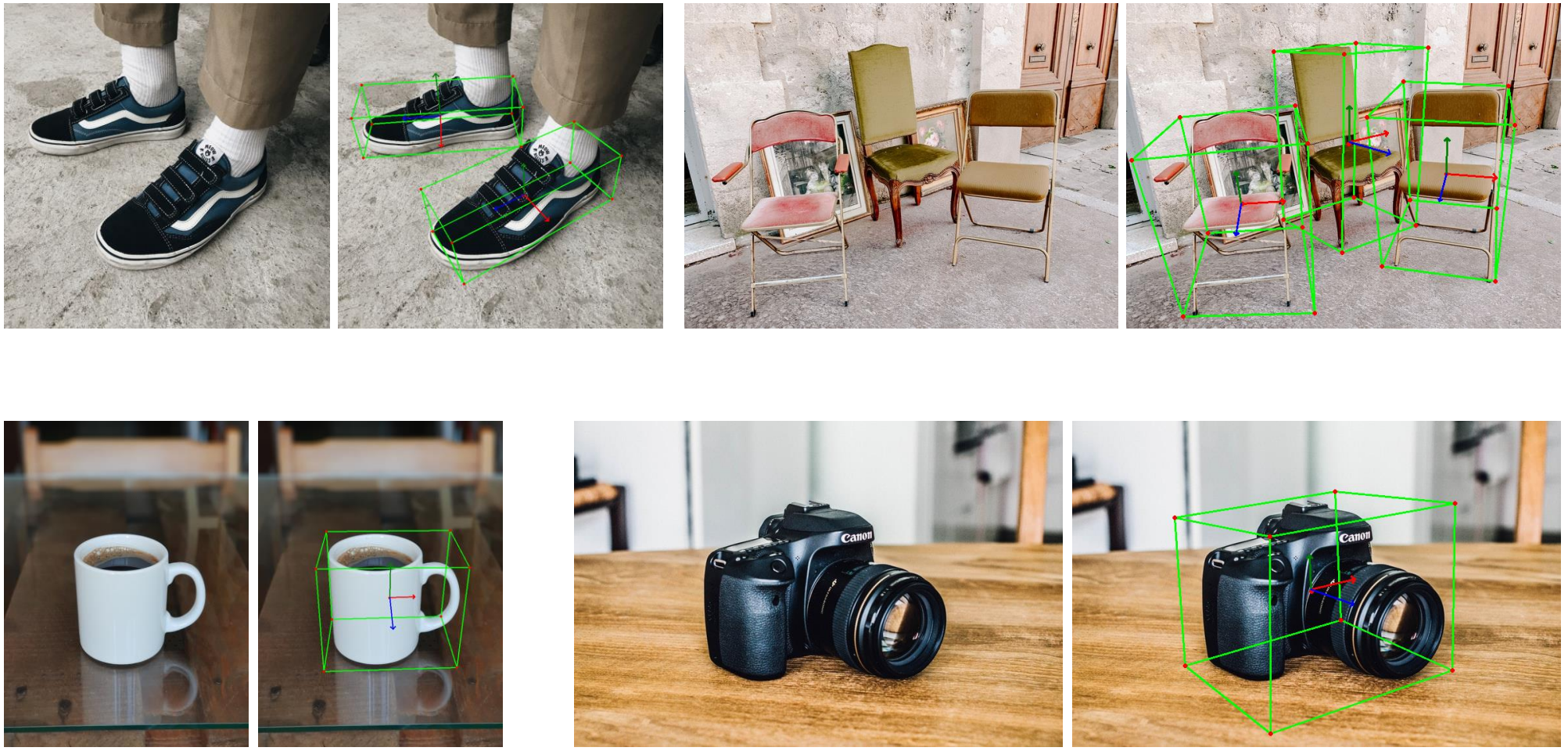


0. WRIST
1. THUMB_CMC
2. THUMB_MCP
3. THUMB_IP
4. THUMB_TIP
5. INDEX_FINGER_MCP
6. INDEX_FINGER_PIP
7. INDEX_FINGER_DIP
8. INDEX_FINGER_TIP
9. MIDDLE_FINGER_MCP
10. MIDDLE_FINGER_PIP
11. MIDDLE_FINGER_DIP
12. MIDDLE_FINGER_TIP
13. RING_FINGER_MCP
14. RING_FINGER_PIP
15. RING_FINGER_DIP
16. RING_FINGER_TIP
17. PINKY_MCP
18. PINKY_PIP
19. PINKY_DIP
20. PINKY_TIP

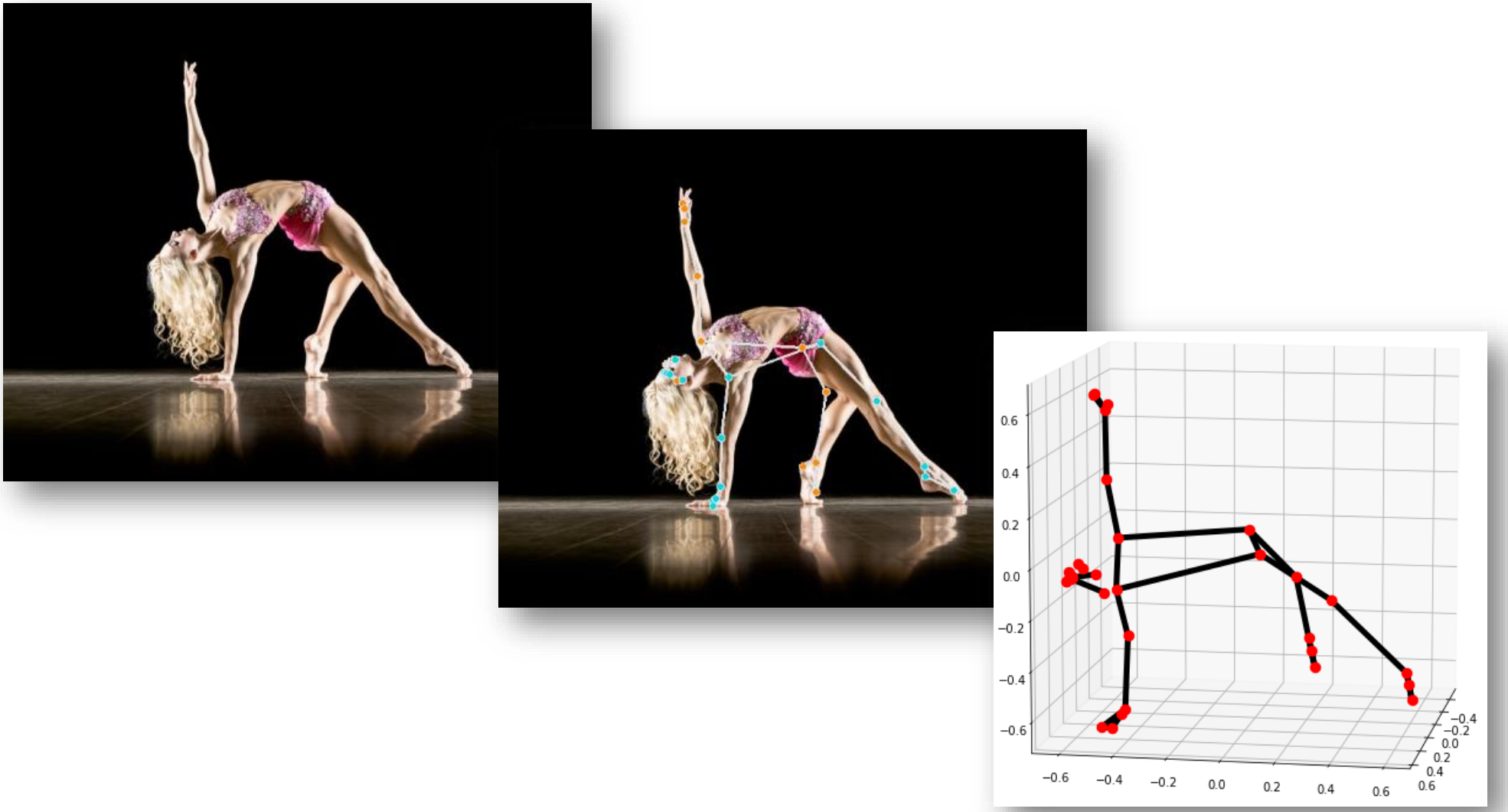
資料來源：https://google.github.io/mediapipe/solutions/hands.html

MediaPipe 面部網點、手指肢體

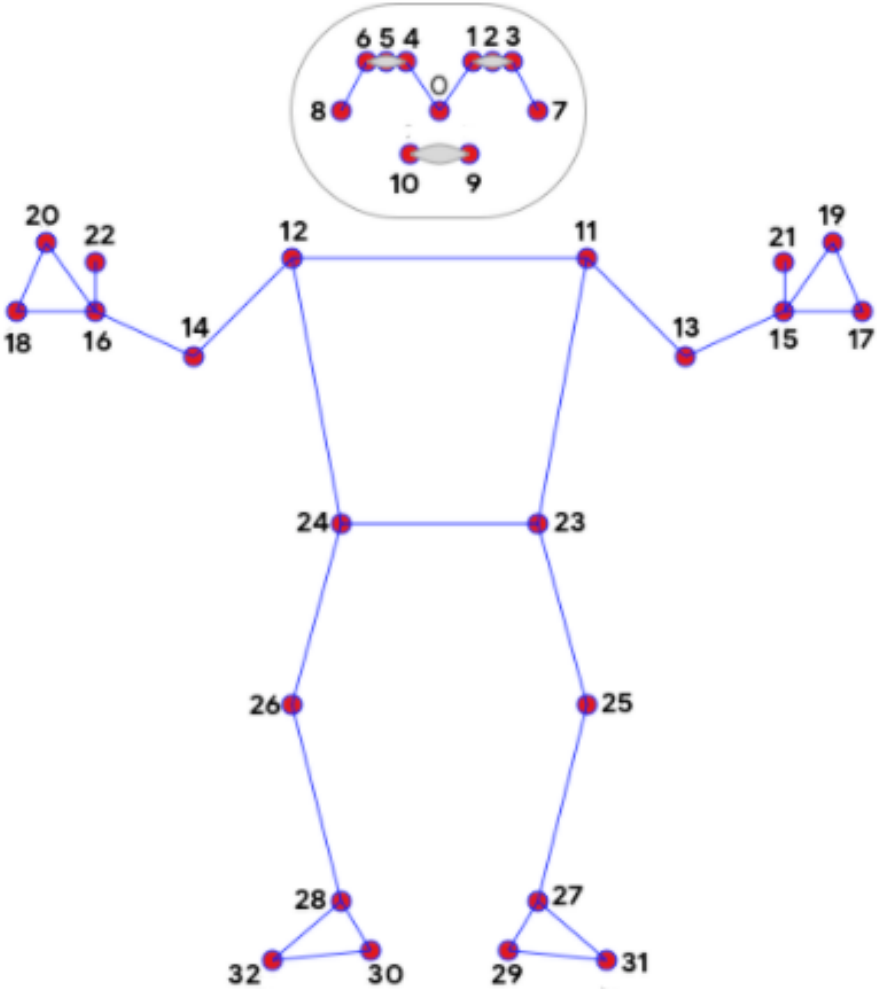
MediaPipe 物件盒偵測



MediaPipe 人體姿態（骨架）



MediaPipe 人體關鍵點定義



0. nose
1. left_eye_inner
2. left_eye
3. left_eye_outer
4. right_eye_inner
5. right_eye
6. right_eye_outer
7. left_ear
8. right_ear
9. mouth_left
10. mouth_right
11. left_shoulder
12. right_shoulder
13. left_elbow
14. right_elbow
15. left_wrist
16. right_wrist
17. left_pinky
18. right_pinky
19. left_index
20. right_index
21. left_thumb
22. right_thumb
23. left_hip
24. right_hip
25. left_knee
26. right_knee
27. left_ankle
28. right_ankle
29. left_heel
30. right_heel
31. left_foot_index
32. right_foot_index

資料來源：https://google.github.io/mediapipe/solutions/pose.html

MediaPipe 人體分割（去背）

人體（遮罩）、背景語義分割

參考文獻

- 許哲豪，【OpenVINO™教學】土炮體感控制系統

https://omnixri.blogspot.com/2019/09/openvino_20.html

- 李盛安，【AI進教室】透過電腦視覺掌握教室學習狀態

https://makerpro.cc/2021/06/use-computer-vision-to-observe-the-teaching-situation-of-computer-laboratory/

- CMU OpenPose Github

https://github.com/CMU-Perceptual-Computing-Lab/openpose

- Google MediaPipe

https://google.github.io/mediapipe/